HEPATITIS C VIRUS ASIALOGLYCOPROTEINS

Related Applications

This application is a continuation-in-part of copending U.S. Serial No. 07/758,880, filed September 13, 1991, which is a continuation-in-part of U.S. Serial No. 07/611,419, filed November 8, 1990, now abandoned, the discosures of which are incorporated herein by reference.

Description

Technical Field

This invention relates to the general fields of recombinant protein expression and virology. More particularly, the invention relates to glycoproteins useful for diagnosis, treatment, and prophylaxis of Hepatitis C virus (HCV) infection, and methods for producing such glycoproteins.

Background of the Invention

Non-A, Non-B hepatitis (NANBH) is a transmissible disease (or family of diseases) that is believed to be virally induced, and is distinguishable from other forms of virus-associated liver disease, such as those caused by hepatitis A virus (HAV), hepatitis B virus (HBV), delta hepatitis virus (HDV), cytomegalovirus (CMV) or Epstein-Barr

virus (EBV). Epidemiologic evidence suggests that there may be three types of NANBH: the water-borne epidemic type; the blood or needle associated type; and the sporadically occurring community acquired type. The number of causative agents is unknown. However, a new viral species, hepatitis C virus (HCV) has recently been identified as the primary (if not only) cause of blood-borne NANBH (BB-NANBH). See for example PCT WO89/046699 and U.S. patent application Serial No. 07/355,002, filed 18 May 1989. Hepatitis C appears to be the major form of transfusion-associated hepatitis in a number of countries or regions, including the United States, Europe, and Japan. There is also evidence implicating HCV in induction of hepatocellular carcinoma. Thus, a need exists for an effective method for preventing and treating HCV infection.

The demand for sensitive, specific methods for screening and identifying carriers of HCV and HCV-contaminated blood or blood products is significant. Post-transfusion hepatitis (PTH) occurs in approximately 10% of transfused patients, and HCV accounts for up to 90% of these cases. The major problem in this disease is the frequent progression to chronic liver damage (25-55%).

Patient care, as well as the prevention of transmission of HCV by blood and blood products or by close personal contact, requires reliable diagnostic and prognostic tools to detect nucleic acids, antigens and antibodies related to HCV. In addition, there is also a need for effective vaccines and immunotherapeutic therapeutic agents for the prevention and/or treatment of the disease.

HCV appears in the blood of infected individuals at very low rates relative to other infectious viruses, which makes the virus very difficult to detect. The low viral burden is probably the primary reason that the causative agent of NANB hepatitis went so long undetected. Even though it has now been cloned, HCV still proves difficult to culture and propagate. Accordingly, there is a strong need for recombinant means of producing diagnostic/therapeutic/prophylactic HCV proteins.

Disclosure of the Invention

It has been found that two HCV proteins, E1 and E2, appear to be membrane associated asialoglycoproteins when expressed in recombinant systems. This is surprising because glycoproteins do not usually remain in mannose-terminated form in mammals, but are further modified with other carbohydrates: the mannose-terminated form is typically only transient. In the case of E1 and E2 (as expressed in our systems), the asialoglycoprotein appears to be the final form. E1 (envelope protein 1) is a glycoprotein having a molecular weight of about 35 kD which is translated from the predicted E1 region of the HCV genome. E2 (envelope protein 2) is a glycoprotein having a molecular weight of about 72 kD which is translated from the predicted NS1 (non-structural protein 1) region of the HCV genome, based on the flaviviral model of HCV. As viral glycoproteins are often highly immunogenic, E1 and E2 are prime candidates for use in immunoassays and therapeutic/prophylactic vaccines.

The discovery that E1 and E2 are not sialylated is significant. The particular form of a protein often dictates which cells may serve as suitable hosts for recombinant expression. Prokaryotes such as *E. coli* do not glycosylate proteins, and are generally not suitable for production of glycoproteins for use as antigens because glycosylation is often important for full antigenicity, solubility, and stability of the protein. Lower eukaryotes such as yeast and fungi glycosylate proteins, but are generally unable to add terminal sialic acid residues to the carbohydrate complexes. Thus, yeast-derived proteins may be antigenically distinct from their natural (non-recombinant) counterparts. Expression in mammalian cells is preferred for applications in which the antigenicity of the product is important, as the glycosylation of the recombinant protein should closely resemble that of the wild viral proteins.

New evidence indicates that the HCV virus may gain entry to host cells during infection through either the asialoglycoprotein receptor found on hepatocytes, or through the mannose receptor found on hepatic endothelial cells and macrophages (particularly Kupffer cells). Surprisingly, it has been found that the bulk of natural E1 and E2

do not contain terminal sialic acid residues, but are only core-glycosylated. A small fraction additionally contains terminal N-acetylglucosamine. Accordingly, it is an object of the present invention to provide HCV envelope glycoproteins lacking all or substantially all terminal sialic acid residues.

Another aspect of the invention is a method for producing asialo-E1 or E2, under conditions inhibiting addition of terminal sialic acid, *e.g.*, by expression in yeast or by expression in mammalian cells using antibiotics to facilitate secretion or release.

Another aspect of the invention is a method for purifying E1 or E2 by affinity to lectins which bind terminal mannose residues or terminal N-acetylglucosamine residues.

Another aspect of the invention is an immunogenic composition comprising a recombinant asialoglycoprotein selected from the group consisting of HCV E1 and E2 in combination with a pharmaceutically acceptable vehicle. One may optionally include an immunological adjuvant, if desired.

Another aspect of the invention is an immunoassay reagent, comprising a recombinant asialoglycoprotein selected from the group consisting of HCV E1 and E2 in combination with a suitable support. Another immunoassay reagent of the invention comprises a recombinant asialoglycoprotein selected from the group consisting of HCV E1 and E2 in combination with a suitable detectable label.

Another aspect of the invention concerns dimers and higher-order aggregates of E1 and/or E2. One species of the invention is an E2 complex. Another species of the invention is an E1:E2 heterodimer.

Another aspect of the invention is an HCV vaccine composition comprising E1:E2 aggregates and a pharmaceutically acceptable carrier.

Another aspect of the invention is a method for purifying E1:E2 complexes.

Modes of Carrying Out The Invention

A. Definitions

The term "asialoglycoprotein" refers to a glycosylated protein which is substantially free of sialic acid moieties. Asialoglycoproteins may be prepared recombinantly, or by purification from cell culture or natural sources. Presently preferred asialoglycoproteins are derived from HCV, preferably the glycoproteins E1 and E2, most preferably recombinant E1 and E2 (rE1 and rE2). A protein is "substantially free" of sialic acid within the scope of this definition if the amount of sialic acid residues does not substantially interfere with binding of the glycoprotein to mannose-binding proteins such as GNA. This degree of sialylation will generally be obtained where less than about 40% of the total N-linked carbohydrate is sialic acid, more preferably less than about 30%, more preferably less than about 20%, more preferably less than about 10%, more preferably less than about 5%, and most preferably less than about 2%.

The term "E1" as used herein refers to a protein or polypeptide expressed within the first 400 amino acids of an HCV polyprotein, sometimes referred to as the E or S protein. In its natural form it is a 35 kD glycoprotein which is found strongly membrane-associated. In most natural HCV strains, the E1 protein is encoded in the viral polyprotein following the C (core) protein. The E1 protein extends from approximately amino acid 192 to about aa383 of the full-length polyprotein. The term "E1" as used herein also includes analogs and truncated mutants which are immunologically crossreactive with natural E1.

The term "E2" as used herein refers to a protein or polypeptide expressed within the first 900 amino acids of an HCV polyprotein, sometimes referred to as the NS1 protein. In its natural form it is a 72 kD glycoprotein which is found strongly membrane-associated. In most natural HCV strains, the E2 protein follows the E1 protein. The E2 protein extends from approximately aa384 to about aa820. The term "E2" as used herein also includes analogs and truncated mutants which are immunologically crossreactive with natural E2.

The term "aggregate" as used herein refers to a complex of E1 and/or E2 containing more than one E1 or E2 monomer. E1:E1 dimers, E2:E2 dimers, and E1:E2 heterodimers are all "aggregates" within the scope of this definition. Compositions of the invention may also include larger aggregates, and may have molecular weights in excess of 800 kD.

The term "particle" as used herein refers to an E1, E2, or E1/E2 aggregate visible by electron microscopy and having a dimension of at least 20 nm. Preferred particles are those having a roughly spherical appearance and a diameter of approximately 40 nm by electron microscopy.

The term "purified" as applied to proteins herein refers to a composition wherein the desired protein comprises at least 35% of the total protein component in the composition. The desired protein preferably comprises at least 40%, more preferably at least about 50%, more preferably at least about 60%, still more preferably at least about 70%, even more preferably at least about 80%, even more preferably at least about 90%, and most preferably at least about 95% of the total protein component. The composition may contain other compounds such as carbohydrates, salts, lipids, solvents, and the like, without affecting the determination of percentage purity as used herein. An "isolated" HCV asialoglycoprotein intends an HCV asialoglycoprotein composition which is at least 35% pure.

"Mannose-binding protein" as used herein intends a lectin or other protein which specifically binds to proteins having mannose-terminated glycosylation (*e.g.*, asialoglycoproteins), for example, mannose-binding lectins, antibodies specific for mannose-terminated glycosylation, mannose receptor protein (R.A.B. Ezekowitz *et al.*, J Exp Med (1990) 176:1785-94), asialoglycoprotein receptor proteins (H. Kurata *et al.*, J Biol Chem (1990) 265:11295-98), serum mannose-binding protein (I. Schuffenecker *et al.*, Cytogenet Cell Genet (1991) 56:99-102; K. Sastry *et al.*, J Immunol (1991) 147:692-97), serum asialoglycoprotein-binding protein, and the like. Mannose-binding lectins include, for

example, GNA, Concanavalin A (ConA), and other lectins with similar binding properties.

The term "GNA lectin" refers to *Galanthus nivalus* agglutinin, a commercially available lectin which binds to mannose-terminated glycoproteins.

A "recombinant" glycoprotein as used herein is a glycoprotein expressed from a recombinant polynucleotide, in which the structural gene encoding the glycoprotein is expressed under the control of regulatory sequences not naturally adjacent to the structural gene, or in which the structural gene is modified. For example, one may form a vector in which the E1 structural gene is placed under control of a functional fragment of the yeast glyceraldehyde-3-phosphate dehydrogenase (GAPDH) promoter. A presently preferred promoter for use in yeast is the hybrid ADH2/GAP promoter described in U.S. Pat. No. 4,880,734 (incorporated herein by reference), which employs a fragment of the GAPDH promoter in combination with the upstream activation sequence derived from alcohol dehydrogenase 2. Modifications of the structural gene may include substitution of different codons with degenerate codons (*e.g.*, to utilize host-preferred codons, eliminate or generate restriction enzyme cleavage sites, to control hairpin formation, *etc.*), and substitution, insertion or deletion of a limited number of codons encoding different amino acids (preferably no more than about 10%, more preferably less than about 5% by number of the natural amino acid sequence should be altered), and the like. Similarly, a "recombinant" receptor refers to a receptor protein expressed from a recombinant polynucleotide, in which the structural gene encoding the receptor is expressed under the control of regulatory sequences not naturally adjacent to the structural gene, or in which the structural gene is modified.

The term "isolated polypeptide" refers to a polypeptide which is substantially free of other HCV viral components, particularly polynucleotides. A polypeptide composition is "substantially free" of another component if the weight of the polypeptide in the composition is at least 70% of the weight of the polypeptide and other component combined, more preferably at least about 80%, still more preferably about 90%, and most

preferably 95% or greater. For example, a composition containing 100 μg/ml E1 and only 3 μg/ml other HCV components (*e.g.*, DNA, lipids, *etc.*) is substantially free of "other HCV viral components", and thus is a composition of an isolated polypeptide within the scope of this definition.

The term "secretion leader" refers to a polypeptide which, when encoded at the N-terminus of a protein, causes the protein to be secreted into the host cell's culture medium following translation. The secretion leader will generally be derived from the host cell employed. For example, suitable secretion leaders for use in yeast include the *Saccharomyces cerevisiae* α-factor leader (see U.S. Pat. No. 4,870,008, incorporated herein by reference).

The term "lower eukaryote" refers to host cells such as yeast, fungi, and the like. Lower eukaryotes are generally (but not necessarily) unicellular. Preferred lower eukaryotes are yeasts, particularly species within *Saccharomyces*, *Schizosaccharomyces*, *Kluveromyces*, *Pichia*, *Hansenula*, and the like. *Saccharomyces cerevisiae*, *S. carlsbergensis* and *K. lactis* are the most commonly used yeast hosts, and are convenient fungal hosts.

The term "higher eukaryote" refers to host cells derived from higher animals, such as mammals, reptiles, insects, and the like. Presently preferred higher eukaryote host cells are derived from Chinese hamster (*e.g.*, CHO), monkey (*e.g.*, COS cells), human, and insect (*e.g.*, *Spodoptera frugiperda*). The host cells may be provided in suspension or flask cultures, tissue cultures, organ cultures, and the like.

The term "calcium modulator" refers to a compound capable of sequestering or binding calcium ions within the endoplasmic reticulum, or affects calcium ion concentration within the ER by its effect on calcium regulatory proteins (*e.g.*, calcium channel proteins, calcium pumps, *etc.*). Suitable calcium modulators include, for example thapsigargin, EGTA (ethylene glycol bis[β-aminoethyl ether] N,N,N',N'-tetraacetic acid). The presently preferred modulator is thapsigargin (see *e.g.*, O. Thastrup *et al*, Proc Nat Acad Sci USA (1990) 87:2466-70).

The term "immunogenic" refers to the ability of a substance to cause a humoral and/or cellular immune response, whether alone or when linked to a carrier, in the presence or absence of an adjuvant. "Neutralization" refers to an immune response that blocks the infectivity, either partially or fully, of an infectious agent. A "vaccine" is an immunogenic composition capable of eliciting protection against HCV, whether partial or complete, useful for treatment of an individual.

The term "biological liquid" refers to a fluid obtained from an organism, such as serum, plasma, saliva, gastric secretions, mucus, and the like. In general, a biological liquid will be screened for the presence of HCV particles. Some biological fluids are used as a source of other products, such as clotting factors (*e.g.*, Factor VIII:C), serum albumin, growth hormone, and the like. In such cases, it is important that the source biological fluid be free of contamination by virus such as HCV.

B. General Method

The E1 region of the HCV genome is described in EP 388,232 as region "E", while E2 is described as "NS1." The E1 region comprises approximately amino acids 192-383 in the full-length viral polyprotein. The E2 region comprises approximately amino acids 384-820. The complete sequences of prototypes of these proteins (strain HCV-1) are available in the art (see EP 388,232), as are general methods for cloning and expressing the proteins. Both E1 and E2 may be expressed from a polynucleotide encoding the first 850-900 amino acids of the HCV polyprotein: post-translational processing in most eukaryotic host cells cleaves the initial polyprotein into C, E1, and E2. One may truncate the 5' end of the coding region to reduce the amount of C protein produced.

Expression of asialoglycoproteins may be achieved by a number of methods. For example, one may obtain expression in lower eukaryotes (such as yeast) which do not normally add sialic acid residues to glycosylated proteins. In yeast expression systems, it is presently preferred to employ a secretion leader such as the *S. cerevisiae* α-factor leader, so that the protein is expressed into the culture medium following

translation. It is also presently preferred to employ glycosylation-deficient mutants such as pmr1, as these mutants supply only core glycosylation, and often secrete heterologous proteins with higher efficiency (H.K. Rudolph *et al*, Cell (1989) 58:133-45). Alternatively, one may employ other species of yeast, such as *Pichia pastoris*, which express glycoproteins containing 8-9 mannose residues in a pattern believed to resemble the core glycosylation pattern observed in mammals and *S. cerevisiae*.

Alternatively, one may arrange expression in mammalian cells, and block terminal glycosylation (addition of sialic acid). Recombinant constructs will preferably include a secretion signal to insure that the protein is directed toward the endoplasmic reticulum. Transport to the golgi appears to be blocked by E1 and E2 themselves: high-level expression of E1 or E2 in mammalian cells appears to arrest secretion of all cellular proteins at the endoplasmic reticulum or *cis* golgi. One may additionally employ a glycosylation defective mutant. See for example, P. Stanley, Ann Rev Genet (1984) 18:525-52. In the event a glycosylation or transport mutant expresses E1 or E2 with sialylation, the terminal sialic acid residues may be removed by treatment with neuraminidase.

Yield should be further increased by use of a calcium modulator to obtain release of protein from within the endoplasmic reticulum. Suitable modulators include thapsigargin, EGTA, and A23817 (see *e.g.*, O. Thastrup *et al*, Proc Nat Acad Sci USA (1990) 87:2466-70). For example, one may express a large amount of E1 or E2 intracellularly in mammalian cells (*e.g.*, CHO, COS, HeLa cells, and the like) by transfection with a recombinant vaccinia virus vector. After allowing time for protein expression and accumulation in the endoplasmic reticulum, the cells are exposed to a calcium modulator in concentration large enough to cause release of the ER contents. The protein is then recovered from the culture medium, which is replaced for the next cycle.

Additionally, it may be advantageous to express a truncated form of the envelope protein. Both E1 and E2 appear to have a highly hydrophobic domain, which apparently anchors the protein within the endoplasmic reticulum and prevents efficient release. Thus, one may wish to delete portions of the sequence found in one or more of

the regions aa170-190, aa260-290 or aa330-380 of E1 (numbering from the beginning of the polyprotein), and aa660-830 of E2 (see for example Figure 20-1 of EP 388,232). It is likely that at least one of these hydrophobic domains forms a transmembrane region which is not essential for antigenicity of the protein, and which may thus be deleted without detrimental effect. The best region to delete may be determined by conducting a small number of deletion experiments within the skill of the ordinary practitioner. Deletion of the hydrophobic 3' end of E2 results in secretion of a portion of the E2 expressed, with sialylation of the secreted protein.

One may use any of a variety of vectors to obtain expression. Lower eukaryotes such as yeast are typically transformed with plasmids using the calcium phosphate precipitation method, or are transfected with a recombinant virus. The vectors may replicate within the host cell independently, or may integrate into the host cell genome. Higher eukaryotes may be transformed with plasmids, but are typically infected with a recombinant virus, for example a recombinant vaccinia virus. Vaccinia is particularly preferred, as infection with vaccinia halts expression of host cell proteins. Presently preferred host cells include HeLa and plasmacytoma cell lines. In the present system, this means that E1 and E2 accumulate as the major glycosylated species in the host ER. As the rE1 and rE2 will be the predominant glycoproteins which are mannose-terminated, they may easily be purified from the cells by using lectins such as *Galanthus nivalus* agglutinin (GNA) which bind terminal mannose residues.

Proteins which are naturally expressed as mannose-terminated glycoproteins are relatively rare in mammalian physiology. In most cases, a mammalian glycoprotein is mannose-terminated only as a transient intermediate in the glycosylation pathway. The fact that HCV envelope proteins, expressed recombinantly, contain mannose-terminated glycosylation or (to a lesser degree) N-acetylglucosamine means that HCV proteins and whole virions may be separated and partially purified from endogenous proteins using lectins specific for terminal mannose or N-acetylglucosamine. The recombinant proteins appear authentic, and are believed essentially identical to the envelope proteins found in

the mature, free virion, or to a form of cell-associated envelope protein. Thus, one may employ lectins such as GNA for mannose-terminated proteins, and WGA (wheat germ agglutinin) and its equivalents for N-acetylglucosamine-terminated proteins. One may employ lectins bound to a solid phase (*e.g.*, a lectin-Sepharose® column) to separate E1 and E2 from cell culture supernatants and other fluids, *e.g.*, for purification during the production of antigens for vaccine or immunoassay use.

Alternatively, one may provide a suitable lectin to isolate E1, E2, or HCV virions from fluid or tissue samples from subjects suspected of HCV infection. As mannose-terminated glycoproteins are relatively rare, such a procedure should serve to purify the proteins present in a sample, substantially reducing the background. Following binding to lectin, the HCV protein may be detected using anti-HCV antibodies. If whole virions are present, one may alternatively detect HCV nucleic acids using PCR techniques or other nucleic acid amplification methods directed toward conserved regions of the HCV genome (for example, the 5' non-coding region). This method permits isolation and characterization of differing strains of HCV without regard for antigenic drift or variation, *e.g.*, in cases where a new strain is not immunologically crossreactive with the strain used for preparing antibodies. There are many other ways to take advantage of the unique recognition of mannose-terminated glycoproteins by particular lectins. For example, one may incubate samples suspected of containing HCV virions or proteins with biotin or avidin-labeled lectins, and precipitate the protein-lectin complex using avidin or biotin. One may also use lectin affinity for HCV proteins to target compounds to virions for therapeutic use, for example by conjugating an antiviral compound to GNA. Alternatively, one may use suitable lectins to remove mannose-terminated glycoproteins from serum or plasma fractions, thus reducing or eliminating the risk of HCV contamination.

It is presently preferred to isolate E1 and/or E2 asialoglycoproteins from crude cell lysates by incubation with an immobilized mannose-binding protein, particularly a lectin such as ConA or GNA. Cells are lysed, *e.g.*, by mechanical disruption in a hypo-

tonic buffer followed by centrifugation to prepare a post-nuclear lysate, and further centrifuged to obtain a crude microsomal membrane fraction. The crude membrane fraction is subsequently solubilized in a buffer containing a detergent, such as Triton X-100, NP40, or the like. This detergent extract is further clarified of insoluble particulates by centrifugation, and the resulting clarified lysate incubated in a chromatography column comprising an immobilized mannose-binding protein, preferably GNA bound to a solid support such as agarose or Sepharose® for a period of time sufficient for binding, typically 16 to 20 hours. The suspension is then applied to the column until E1/E2 begins to appear in the eluent, then incubated in the column for a period of time sufficient for binding, typically about 12-24 hours. The bound material is then washed with additional buffer containing detergent (*e.g.*, Triton X-100, NP40, or the like), and eluted with mannose to provide purified asialoglycoprotein. On elution, it is preferred to elute only until protein begins to appear in the eluate, at which point elution is halted and the column permitted to equilibrate for 2-3 hours before proceeding with elution of the protein. This is believed to allow sufficient time for the slow off-rate expected of large protein aggregates. In cases wherein E1 and E2 are expressed together in native form (*i.e.*, without truncation of the membrane-binding domain), a substantial fraction of the asialoglycoproteins appear as E1:E2 aggregates. When examined by electron microscopy, a significant portion of these aggregates appear as roughly spherical particles having a diameter of about 40 nm, which is the size expected for intact virus. These particles appear to be self-assembling subviral particles. These aggregates are expected to exhibit a quaternary structure very similar to the structure of authentic HCV virion particles, and thus are expected to serve as highly immunogenic vaccines.

The E1/E2 complexes may be further purified by gel chromatography on a basic medium, for example, Fractogel-DEAE or DEAE-Sepharose®. Using Fractogel-DEAE gel chromatography, one may obtain E1/E2 complexes of approximately 60-80% purity. One may further purify E1 by treatment with lysine protease, because E1 has 0-1

Lys residues. Treatment of the complex with lysine protease destroys E2, and permits facile separation of E1.

The tissue specificity of HCV, in combination with the observation that HCV envelope glycoproteins are mannose-terminated, suggests that the virus employs the mannose receptor or the asialoglycoprotein receptor (ASGR) in order to gain entry into host cells. Mannose receptors are found on macrophages and hepatic sinusoidal cells, while the ASGR is found on parenchymal hepatocytes. Thus, it should be possible to culture HCV by employing host cells which express one or both of these receptors. One may either employ primary cell cultures which naturally express the receptor, using conditions under which the receptor is maintained, or one may transfect another cell line such as HeLa, CHO, COS, and the like, with a vector providing for expression of the receptor. Cloning of the mannose receptor and its transfection and expression in fibroblasts has been demonstrated by M.E. Taylor *et al*, J Biol Chem (1990) 265:12156-62. Cloning and sequencing of the ASGR was described by K. Drickamer *et al*, J Biol Chem (1984) 259:770-78 and M. Spiess *et al*, Proc Nat Acad Sci USA (1985) 82:6465-69; transfection and expression of functional ASGR in rat HTC cells was described by M. McPhaul and P. Berg, Proc Nat Acad Sci USA (1986) 83:8863-67 and M. McPhaul and P. Berg, Mol Cell Biol (1987) 7:1841-47. Thus, it is possible to transfect one or both receptors into suitable cell lines, such as CHO, COS, HeLa, and the like, and to use the resulting cells as hosts for propagation of HCV in culture. Serial passaging of HCV in such cultures should result in development of attenuated strains suitable for use as live vaccines. It is presently preferred to employ an immortalized cell line transfected with one or both recombinant receptors.

Immunogenic compositions can be prepared according to methods known in the art. The present compositions comprise an immunogenic amount of a polypeptide, *e.g.*, E1, E2, or E1/E2 particle compositions, usually combined with a pharmaceutically acceptable carrier, preferably further comprising an adjuvant. If a "cocktail" is desired, a combination of HCV polypeptides, such as, for example, E1 plus E2 antigens, can be

mixed together for heightened efficacy. The virus-like particles of E1/E2 aggregates are expected to provide a particularly useful vaccine antigen. Immunogenic compositions may be administered to animals to induce production of antibodies, either to provide a source of antibodies or to induce protective immunity in the animal.

Pharmaceutically acceptable carriers include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers; and inactive virus particles. Such carriers are well known to those of ordinary skill in the art.

Preferred adjuvants to enhance effectiveness of the composition include, but are not limited to: aluminum hydroxide (alum), N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP) as found in U.S. Patent No. 4,606,918, N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-hydroxyphosphoryloxy)-ethylamine (MTP-PE) and RIBI, which contains three components extracted from bacteria, monophosphoryl lipid A, trehalose dimycolate, and cell wall skeleton (MPL+TDM+CWS) in a 2% squalene/Tween® 80 emulsion. Additionally, adjuvants such as Stimulon (Cambridge Bioscience, Worcester, MA) may be used. Further, Complete Freund's Adjuvant (CFA) and Incomplete Freund's Adjuvant (IFA) may be used for non-human applications and research purposes.

The immunogenic compositions typically will contain pharmaceutically acceptable vehicles, such as water, saline, glycerol, ethanol, *etc.* Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be included in such vehicles.

Typically, the immunogenic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect.

Immunogenic compositions used as vaccines comprise an immunologically effective amount of the HCV polypeptide, as well as any other of the above-mentioned components, as needed. "Immunologically effective amount", means that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment, as defined above. This amount varies depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated (*e.g.*, nonhuman primate, primate, *etc.*), the capacity of the individual's immune system to synthesize antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, the strain of infecting HCV, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

The self-assembling E1/E2 aggregates may also serve as vaccine carriers to present heterologous (non-HCV) haptens, in the same manner as Hepatitis B surface antigen (See European Patent Application 174,444). In this use, the E1/E2 aggregates provide an immunogenic carrier capable of stimulating an immune response to haptens or antigens conjugated to the aggregate. The antigen may be conjugated either by conventional chemical methods, or may be cloned into the gene encoding E1 and/or E2 at a location corresponding to a hydrophilic region of the protein.

The immunogenic compositions are conventionally administered parenterally, typically by injection, for example, subcutaneously or intramuscularly. Additional formulations suitable for other modes of administration include oral formulations and suppositories. Dosage treatment may be a single dose schedule or a multiple dose schedule. The vaccine may be administered in conjunction with other immunoregulatory agents.

C. Examples

The examples presented below are provided as a further guide to the practitioner of ordinary skill in the art, and are not to be construed as limiting the invention in any way.

Example 1

(Cloning and Expression)

(A) Vectors were constructed from plasmids containing the 5 ' portion of the HCV genome, as described in EP 318,216 and EP 388,232. Cassette HCV(S/B) contains a StuI-BglII DNA fragment encoding the 5 ' end of the polyprotein from Met_1 up to Leu_{906}, beginning at nucleotide -63 relative to Met_1. This includes the core protein (C), the E1 protein (also sometimes referred to as S), the E2 protein (also referred to as NS1), and a 5 ' portion of the NS2a region. Upon expression of the construct, the individual C, E1 and E2 proteins are produced by proteolytic processing.

Cassette HCV(A/B) contains a ApaLI-BglII DNA fragment encoding the 5 ' end of the polyprotein from Met_1 up to Leu_{906}, beginning at nucleotide -6 relative to Met_1. This includes the core protein (C), the E1 protein (also sometimes referred to as S), the E2 protein (also referred to as NS1), and a 5 ' portion of the NS2a region. Upon expression of the construct, the individual C, E1 and E2 proteins are produced by proteolytic processing.

Cassette C-E1(S/B) (a StuI-BamHI portion) contains the 5 ' end from Met_1 up to Ile_{340} (a BamHI site in the gene). Expression of this cassette results in expression of C and a somewhat truncated E1 (E1 '). The portion truncated from the 3 ' end is a hydrophobic region believed to serve as a translocation signal.

Cassette NS1(B/B) (a BamHI-BglII portion) contains a small 3 ' portion of E1 (from Met_{364}), all of E2, and a portion of NS2a (to Leu_{906}). In this construct, the E1 fragment serves as a translocation signal.

Cassette TPA-NS1 employs a human tissue plasminogen activator (tPA) leader as a translocation signal instead of the 3 ' portion of E1. The cassette contains a truncated form of E2, from Gly_{406} to Glu_{661}, in which the hydrophobic 3 ' end is deleted.

Each cassette was inserted into the vector pGEM3Z (Promega) with and without a synthetic β-globin 5 ' non-coding sequence for transcription and translation using T7 and rabbit reticulocyte expression *in vitro*. Recombinant vaccinia virus (rVV)

vectors were prepared by inserting the cassettes into the plasmid pSC11 (obtained from Dr. B. Moss, NIH) followed by recombination with vaccinia virus, as described by Charkrabarty *et al*, Mol Cell Biol (1985) 5:3403-09.

(B) An alternate expression vector was constructed by inserting HCV(A/B) between the StuI and SpeI sites of pSC59 (obtained from Dr. B. Moss, NIH) followed by recombination with vaccinia virus, as described by Charkrabarty *et al*, Mol Cell Biol (1985) 5:3403-09.

(C) HeLa S3 cells were collected by centrifugation for 7 minutes at 2000 rpm at room temperature in sterile 500 ml centrifuge bottles (JA-10 rotor). The pellets were resuspended at a final concentration of 2×10^7 cells/ml in additional culture medium (Joklik modified MEM Spinner medium +5% horse serum and Gentamycin) ("spinner medium"). Sonicated crude vv/SC59-HCV virus stock was added at a multiplicity of infection of 8 pfu/cell, and the mixture stirred at 37°C for 30 minutes. The infected cells were then transferred to a spinner flask containing 8 liters spinner medium and incubated for 3 days at 37°C.

The cultured cells were then collected by centrifugation, and the pellets resuspended in buffer (10 mM Tris-HCl, pH 9.0, 15 ml). The cells were then homogenized using a 40 ml Dounce Homogenizer (50 strokes), and the nuclei pelleted by centrifugation (5 minutes, 1600 rpm, 4°C, JA-20 rotor). The nuclear pellets were resuspended in Tris buffer (4 ml), rehomogenized, and pelleted again, pooling all supernatants.

The pooled lysate was divided into 10 ml aliquots and sonicated 3×30 minutes in a cuphorn sonicator at medium power. The sonicated lysate (15 ml) was layered onto 17 ml sucrose cushions (36%) in SW28 centrifuge tubes, and centrifuged at 13,500 rpm for 80 minutes at 4°C to pellet the virus. The virus pellet was resuspended in 1 ml of Tris buffer (1 mM Tris HCl, pH 9.0) and frozen at -80°C.

Example 2

(Comparison of *In Vitro* and *In Vivo* Products)

(A) E1 and E2 were expressed both *in vitro* and *in vivo* and ^{35}S-Met labeled using the vectors described in Example 1 above. BSC-40 and HeLa cells were infected with the rVV vectors for *in vivo* expression. Both the medium and the cell lysates were examined for recombinant proteins. The products were immunoprecipitated using human HCV immune serum, while *in vitro* proteins were analyzed directly. The resulting proteins were analyzed by SDS-PAGE.

The reticulocyte expression system (pGEM3Z with HCV(S/B) or HCV(A/B)) produced C, E1 and E2 proteins having molecular weights of approximately 18 kD, 35 kD, and 72 kD, respectively. Lysates from BSC-40 and HeLa cells transfected with rVV containing HCV(S/B), HCV(A/B) or C-E1(S/B) exhibited the same proteins. Because the reticulocyte system does not provide efficient golgi processing and therefore does not provide sialic acid, the fact that both *in vitro* and *in vivo* products exhibited identical mobilities suggests that the proteins are not sialylated *in vivo*. Only the rVV vector containing TPA-NS1 resulted in any extracellular secretion of E2, which exhibited an altered mobility consistent with sialylation.

(B) HCV(S/B) was expressed *in vitro* and incubated with a panel of biotinylated lectins: GNA, SNA, PNA, WGA, and ConA. Following incubation, the complexes were collected on avidin-acrylic beads, washed, eluted with Laemmli sample buffer, and analyzed by SDS-PAGE. The results showed that E1 and E2 bound to GNA and ConA, which indicates the presence of mannose. GNA binds to terminal mannose groups, while ConA binds to any α-linked mannose. The lack of binding to SNA, PNA, and WGA indicates that none of the proteins contained sialic acid, galactose-N-acetylgalactosamine, or N-acetylglucosamine.

(C) Radiolabeled E1 and E2 were produced in BSC-40 cells by infection with rVV containing HCV(S/B) (vv/SC11-HCV), and immunoprecipitated with human HCV$^+$ immune serum. One half of the immunoprecipitated material was treated overnight with

neuraminidase to remove any sialic acid. Following treatment, the treated and untreated proteins were analyzed by SDS-PAGE. No significant difference in mobility was observed, indicating lack of sialylation *in vivo*.

(D) Radiolabeled E1 and E2 were produced in BSC-40 cells by infection with rVV containing HCV(A/B) (vv/SC59-HCV), and either immunoprecipitated with human HCV^+ serum, or precipitated using biotinylated GNA lectin linked to acrylic beads, using vv/SC11 free of HCV sequences as control. The precipitates were analyzed by SDS-PAGE. The data demonstrated that E1 and E2 were the major species of mannose-terminated proteins in vv/SC59-HCV infected cells. GNA was as efficient as human antisera in precipitating E1 and E2 from cell culture medium. A 25 kD component was observed, but appears to be specific to vaccinia-infected cells.

Example 3

(Purification Using Lectin)

(A) HeLa S3 cells were inoculated with purified high-titer vv/SC59-HCV virus stock at a multiplicity of infection of 5 pfu/cell, and the mixture stirred at 37°C for 30 minutes. The infected cells were then transferred to a spinner flask containing 8 liters spinner medium and incubated for 3 days at 37°C. The cells were collected again by centrifugation and resuspended in hypotonic buffer (20 mM HEPES, 10 mM NaCl, 1 mM $MgCl_2$, 120 ml) on ice. The cells were then homogenized by Dounce Homogenizer (50 strokes), and the nuclei pelleted by centrifugation (5 minutes, 1600 rpm, 4°C, JA-20 rotor). The pellets were pooled, resuspended in 48 ml hypotonic buffer, rehomogenized, recentrifuged, pooled again, and frozen at -80°C.

The frozen supernatants were then thawed, and the microsomal membrane fraction of the post-nuclear lysate isolated by centrifuging for 20 minutes in a JA-20 rotor at 13,500 rpm at 4°C. The supernatant was removed by aspiration.

The pellets were taken up in 96 ml detergent buffer (20 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, 1 mM DDT, 0.5% Triton X-100, pH 7.5) and homogenized

(50 strokes). The product was clarified by centrifugation for 20 minutes at 13,500 rpm, 4°C, and the supernatants collected.

A GNA-agarose column (1 cm × 3 cm, 3 mg GNA/ml beads, 6 ml bed volume, Vector Labs, Burlingame, CA) was pre-equilibrated with detergent buffer. The supernatant sample was applied to the column with recirculation at a flow rate of 1 ml/min for 16-20 hours at 4°C. The column was then washed with detergent buffer.

The purified E1/E2 proteins were eluted with α-D-mannoside (0.9 M in detergent buffer) at a flow rate of 0.5 ml/minute. Elution was halted at the appearance of E1/E2 in the eluent, and the column allowed to reequilibrate for 2-3 hours. Fractions were analyzed by Western blot and silver staining. Peak fractions were pooled and UV-irradiated to inactivate any residual vaccinia virus.

(B) GNA-agarose purified E1 and E2 asialoglycoproteins were sedimented through 20-60% glycerol gradients. The gradients were fractionated and proteins were analyzed by SDS-PAGE and western blotting. Blots were probed with GNA for identification of E1 and E2. The results indicate the presence of a E1:E2 heterodimer which sediments at the expected rate (*i.e.*, a position characteristic of a 110 kD protein). Larger aggregates of HCV envelope proteins also are apparent. E2:E2 homodimers also were apparent. E2 appeared to be over-represented in the larger species relative to E1, although discrete E1:E2 species also were detected. The larger aggregates sedimented significantly faster than the thyroglobulin marker.

(C) GNA-agarose purified E1 and E2 were sedimented through 20-60% glycerol gradients containing 1 mM EDTA. Fractions were analyzed by SDS-PAGE with and without β-mercaptoethanol (βME). Little or no difference in the apparent abundance of E1 and E2 in the presence or absence of βME was observed, indicating the absence of disulfide links between heterodimers.

(D) E1/E2 complexes (approximately 40% pure) were analyzed on a Coulter DM-4 sub-micron particle analyzer. Material in the 20-60 nm range was detected.

(E) E1/E2 complexes (approximately 40% pure) were analyzed by electron microscopy using negative staining with phosphotungstic acid. The electron micrograph revealed the presence of particles having a spherical appearance and a diameter of about 40 nm. E1/E2 complexes were incubated with HCV^+ human immune serum, then analyzed by EM with negative staining. Antibody complexes containing large aggregates and smaller particles were observed.

Example 4

(Chromatographic Purification)

(A) The GNA lectin-purified material prepared as described in Example 3 (0.5-0.8 ml) was diluted 10× with buffer A (20 mM Tris-Cl buffer, pH 8.0, 1 mM EDTA), and applied to a 1.8 × 1.5 cm column of Fractogel EMD DEAE-650 (EM Separations, Gibbstown, New Jersey, cat. no. 16883) equilibrated in buffer A. The protein fraction containing E1/E2 was eluted with the same buffer at a flow rate of 0.2 ml/minute, and 1 ml fractions collected. Fractions containing E1 and E2 (determined by SDS-PAGE) were pooled and stored at -80°C.

(B) The material purified in part (A) above has a purity of 60-80%, as estimated by SDS-PAGE. The identification of the putative E1 and E2 bands was confirmed by N-terminal sequence analysis after using a transfer technique. For the purpose, the fractogel-DEAE purified E1/E2 material was reduced by addition of Laemmli buffer (pH 6.8, 0.06 M Tris-Cl, 2.3% SDS, 10% glycerol, 0.72 M β-mercaptoethanol) and boiled for 3 minutes. The sample was then loaded onto a 10% polyacrylamide gel. After SDS-PAGE, the protein was transferred to a polyvinylidene difluoride (PVDF) 0.2 μm membrane (Bio-Rad Laboratories, Richmond, CA). The respective putative E1 and E2 protein bands were excised from the blot and subjected to N-terminal amino acid analysis, although no special care was taken to prevent amino-terminal blockage during preparation of the material. The first 15 cycles revealed that the E1 sample had a sequence Tyr-Gln-Val-Arg-X-Ser-Thr-Gly-X-Tyr-His-Val-X-Asn-Asp, while the sequence of E2 was Thr-

His-Val-Thr-Gly-X-X-Ala-Gly-His-X-Val-X-Gly-Phe. This amino acid sequence data is in agreement with that expected from the corresponding DNA sequences.

The E1/E2 product purified above by fractogel-DEAE chromatography is believed to be aggregated as evidenced by the fact that a large amount of E1 and E2 coelutes in the void volume region of a gel permeation chromatographic Bio-Sil TSK-4000 SW column. This indicates that under native conditions a significant amount of the E1/E2 complex has a molecular weight of at least 800 kD. E1/E2 material having a molecular weight of about 650 kD was also observed.

Example 5

(Additional Cloning and Expression)

(A) The following cassettes containing 5' portions of the HCV polyprotein were inserted into the vector pGEM4Z (Promega) with and without a synthetic yellow fever virus 5' non-coding sequence and also into recombinant vaccinia virus (rVV) vectors (as described in Example 1A). Cassette C5p-1 contains a fragment encoding the 5' end of the polyprotein from Met_1 to Trp_{1079}, beginning at nucleotide -275 relative to Met_1, with EcoRI linkers on the 5' and 3' ends. Cassette C5p-3 contains an fragment encoding the 5' end of the polyprotein from Met_1 to Trp_{1079}, with EcoRI linkers on the 5' and 3' ends. Both cassettes encode C, E1 and E2 proteins and a 5' portion of the NS2 protein.

(B) The following cassettes containing 5' portions of the HCV polyprotein were inserted into the vector pSC59 followed by recombination with vaccinia virus (described in Example 1B). Cassette HCV(Poly) contains a blunt-ended StuI-BglII fragment encoding the 5' end of the polyprotein from Met_1 to Asp_{966} beginning at nucleotide -65 relative to Met_1. This construct expresses C, E1 and E2 proteins, and a 5' portion of the NS2 protein.

Cassette HCV(5C/SB) contains a blunt-ended StuI-BamHI fragment encoding the 5' end of the polyprotein from Met_1 to Ile_{340} beginning at nucleotide -65 relative to Met_1. This construct expresses the C protein and a truncated E1 protein.

Cassette HCV(6C/SS) contains a SalI(blunted)-EcoRI fragment encoding the 5' end of the polyprotein from Met_1 to Asp_{382} wherein Ser_2 is replaced with Gly_2. This construct expresses the C protein and a truncated E1 protein.

Cassette HCV(E12C/B) contains a blunt-ended ClaI-BglII fragment encoding a portion of the polyprotein from Met_{134} to Asp_{966} inserted into an EcoRI blunted SC59 vector.

Cassette HCV(E1/S) contains a blunt-ended ClaI/SalI fragment encoding a portion of the polyprotein from Met_{134} to Val_{381} inserted into an EcoRI blunted SC59 vector.

(C) HeLa S3 cells were collected by centrifugation for 7 minutes at 2000 rpm at 4°C in sterile 250 ml centrifuge bottles. The pellets were resuspended at a final concentration of 5×10^6 cells/ml in Gey's balanced Salt Solution (GBSS). Sonicated crude vv/SC59-HCV virus stock was added at a multiplicity of infection of 0.5 pfu/cell, and the mixture stirred at 37°C for 1-2 hours. The infected cells were then transferred at a final concentration of 10^6 cells/ml to a spinner flask containing 1 liter culture medium (Joklik MEM + 10% fetal bovine serum + non-essential amino acids, vitamins, pen/strep) and incubated for 3 days at 37°C.

The cultured cells were then collected by centrifugation, and the pellets resuspended in buffer (10 mM Tris-HCl, pH 9.0, 15 ml). The cells were then homogenized using a 40 ml Dounce Homogenizer (50 strokes), and the nuclei pelleted by centrifugation (5 minutes, 1600 rpm, 4°C, JA-20 rotor). The nuclear pellets were resuspended in Tris buffer (4 ml), rehomogenized, and pelleted again, pooling all supernatants.

The pooled lysate was divided into 5 ml aliquots, and 0.1 volume of 2.5 mg/ml trypsin added and incubated at 37°C for 30 minutes. The aliquots were then sonicated 3×30 seconds in a cuphorn sonicator at medium power. The sonicated lysate was used as the crude stock.

Example 6

(Additional Comparison of *In Vitro* and *In Vivo* Products)

(A) E1 and E2 were expressed both *in vitro* and *in vivo* and ^{35}S-Met labeled using the vectors described in Example 5 above and the procedures described in Example 2 above. BSC-40 and HeLa cells were infected with the rVV vectors for *in vivo* expression. Both the medium and the cell lysates were examined for recombinant proteins. The products were immunoprecipitated using human HCV immune serum or rabbit or goat anti-HCV antiserum, while *in vitro* proteins were analyzed directly. The resulting proteins were analyzed by SDS-PAGE and EndoH digestion.

Example 7

(Additional Lectin Purification)

(A) HeLa S3 cells were inoculated with purified high-titer vv/SC59-HCV virus stock (HCV(Poly) or HCV(E12C/B) as described in Example 5 above) at a multiplicity of infection of 1 pfu/cell, and the mixture stirred at 37°C for 1-2 hours. The infected cells were then transferred to spinner flasks containing 1 liter culture medium (see Example 5, supra) and incubated for 2 days at 37°C. A total of 10 liters of cells were collected by centrifugation and resuspended in hypotonic buffer (20 mM HEPES, 10 mM NaCl, 1 mM $MgCl_2$, 120 ml) containing protease inhibitors (PMSF and pepstatin A) on ice. The cells were then homogenized in a 40 ml homogenizer in two batches, pelleted by centrifugation (20 minutes, 12,000 rpm), and re-suspended and re-homogenized. Each pellet was resuspended in approximately 10 ml 25 mM $NaPO_4$ (pH 6.8) in a homogenizer, and an equal volume of 4% Triton X-100 in 100 mM $NaPO_4$ (pH 6.8) added. The pelleted cells were homogenized with 20 strokes, spun at 12,000 rpm for 15 minutes (4 tubes/pellet), and the supernatant saved. The resuspension, Triton addition and centrifugation steps were repeated, and the saved supernatants combined, and frozen at -80°C.

The frozen supernatants were thawed, spun at 12,000 rpm for 15 minutes, combined and held at 4°C. A GNA-agarose column (1 cm × 3 cm, 3 mg GNA/ml

beads, 6 ml bed volume, Vector Labs, Burlingame, CA) was pre-equilibrated with detergent buffer (2% Triton X-100 in 50 mM $NaPO_4$, pH 6.8). The supernatant sample was applied to the column with recirculation at a flow rate of 1 ml/min for 16-20 hours at 4°C. The column was then washed with detergent buffer, followed by 30 ml each of: Buffer A (1M NaCl, 20 mM $NaPO_4$ (pH 6.0), 0.1% Triton X-100); Buffer B (20 mM $NaPO_4$ (pH 6.0), 0.1% Triton X-100); Buffer D (0.2M methyl-α-D-mannopyranoside (mmp), 20 mM $NaPO_4$ (pH 6.0), 0.1% Triton X-100); Buffer E (1M mmp, 20 mM $NaPO_4$ (pH 6.0), 0.1% Triton X-100); Buffer F (1M mmp, 1M NaCl, 20 mM $NaPO_4$ (pH 6.0), 0.1% Triton X-100). Purified E1/E2 proteins come off as eluted material in Buffers D and E, which were collected separately and analyzed by SDS-PAGE.

Example 8

(Additional Chromatographic Purification)

(A) The GNA lectin-purified material prepared as described in Example 7 was applied to a column of S-Sepharose Fast Flow (Pharmacia) equilibrated in buffer B (see Example 7). The column was washed with Buffer B, then eluted with Buffer 1 (0.5M NaCl, 20 mM $NaPO_4$ (pH 6.0), 0.1% Triton X-100) and Buffer 2 (1M NaCl, 20 mM $NaPO_4$ (pH 6.0), 0.1% Triton X-100). Fractions containing E1 and E2 (determined by SDS-PAGE) were pooled and stored at -80°C.